

11-01-00

A

Practitioner's Docket No. 944-003.040

PATENT

## Preliminary Classification:

Proposed Class:

Subclass:

NOTE: "All applicants are requested to include a preliminary classification on newly filed patent applications. The preliminary classification, preferably class and subclass designations, should be identified in the upper right-hand corner of the letter of transmittal accompanying the application papers, for example "Proposed Class 2, subclass 129." M.P.E.P. § 601, 7<sup>th</sup> ed.

## IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

**Box Patent Application**  
**Assistant Commissioner for Patents**  
**Washington, D.C. 20231**

## NEW APPLICATION TRANSMITTAL

Transmitted herewith for filing is the patent application of

Inventor(s): Jari MÄKINEN, Hannu MIKKOLA, Janne VAINIO and Jani ROTOLA-PUKKILA

**WARNING:** 37 C.F.R. § 1.41(a)(1) points out:

"(a) A patent is applied for in the name or names of the actual inventor or inventors.

"(1) The inventorship of a nonprovisional application is that inventorship set forth in the oath or declaration as prescribed by § 1.63, except as provided for in § 1.53(d)(4) and § 1.63(d). If an oath or declaration as prescribed by § 1.63 is not filed during the pendency of a nonprovisional application, the inventorship is that inventorship set forth in the application papers filed pursuant to § 1.53(b), unless a petition under this paragraph accompanied by the fee set forth in § 1.17(l) is filed supplying or changing the name or names of the inventor or inventors."

For (title): **METHOD AND SYSTEM FOR SPEECH FRAME ERROR  
 CONCEALMENT IN SPEECH DECODING**

## CERTIFICATION UNDER 37 C.F.R. § 1.10\*

(Express Mail label number is **mandatory**)  
 (Express Mail certification is optional.)

I hereby certify that this New Application Transmittal and the documents referred to as attached therein are being deposited with the United States Postal Service on this date October 31, 2000, in an envelope as "Express Mail Post Office to Addressee," mailing Label Number EL628639055US, addressed to the: Assistant Commissioner for Patents, Washington, D.C. 20231.

Jennifer A. Hanlon

(type or print name of person mailing paper)

*Jennifer A. Hanlon*  
 Signature of person mailing paper

**WARNING:** Certificate of mailing (first class) or facsimile transmission procedures of 37 C.F.R. § 1.8 cannot be used to obtain a date of mailing or transmission for this correspondence.

**WARNING:** Each paper or fee filed by "Express Mail" **must** have the number of the "Express Mail" mailing label placed thereon prior to mailing. 37 C.F.R. § 1.10(b). Since the filing of correspondence under § 1.10 without the Express Mail mailing label thereon is an oversight that can be avoided by the exercise of reasonable care, requests for waiver of this requirement will **not** be granted on petition." Notice of Oct. 24, 1996, 60 Fed. Reg. 56,439, at 56,442.

10/31/00  
 JCS60 U.S. PRO

JCS60 U.S. PRO  
 09/10/05 40  
 10/31/00

09702540-103100

## 1. Type of Application

This new application is for a(n)

(check one applicable item below)

☒ Original (nonprovisional)

☐ Design

☐ Plant

**WARNING:** Do not use this transmittal for a completion in the U.S. of an International Application under 35 U.S.C. § 37(c)(4), unless the International Application is being filed as a divisional, continuation or continuation-in-part application.

**WARNING:** Do not use this transmittal for the filing of a provisional application.

**NOTE:** If one of the following 3 items apply, then complete and attach ADDED PAGES FOR NEW APPLICATION TRANSMITTAL WHERE BENEFIT OF A PRIOR U.S. APPLICATION CLAIMED and a NOTIFICATION IN PARENT APPLICATION OF THE FILING OF THIS CONTINUATION APPLICATION.

☐ Divisional

☐ Continuation

☐ Continuation-in-part (C-I-P)

## 2. Benefit of Prior U.S. Application(s) (35 U.S.C. §§ 119(e), 120, or 121)

**NOTE:** A nonprovisional application may claim an invention disclosed in one or more prior filed copending nonprovisional applications or copending international applications designating the United States of America. In order for a nonprovisional application to claim the benefit of a prior filed copending nonprovisional application or copending international application designating the United States of America, each prior application must name as an inventor at least one inventor named in the later filed nonprovisional application and disclose the named inventor's invention claimed in at least one claim of the later filed nonprovisional application in the manner provided by the first paragraph of 35 U.S.C. § 112. Each prior application must also be:

(i) An international application entitled to a filing date in accordance with PCT Article 11 and designated the United States of America; or

(ii) Complete as set forth in § 1.51(b); or

(iii) Entitled to a filing date as set forth in § 1.53(b) or § 1.53(d) and include the basic filing fee set forth in § 1.16; or

(iv) Entitled to a filing date as set forth in § 1.53(b) and have paid therein the processing and retention fee set forth in § 1.21(f) within the time period set forth in § 1.53(f).

37 C.F.R. § 1.78(a)(1).

**NOTE:** If the new application being transmitted is a divisional, continuation or a continuation-in-part of a parent case, or where the parent case is an International Application which designated the U.S., or benefit of a prior provisional application is claimed, then check the following item and complete and attach ADDED PAGES FOR NEW APPLICATION TRANSMITTAL WHERE BENEFIT OF PRIOR U.S. APPLICATION(S) CLAIMED.

**WARNING:** If an application claims the benefit of the filing date of an earlier filed application under 35 U.S.C. §§ 120, 121 or 365(c), the 20-year term of that application will be based upon the filing date of the earliest U.S. application that the application makes reference to under 35 U.S.C. §§ 120, 121 or 365(c). (35 U.S.C. § 154(a)(2) does not take into account, for the determination of the patent term, any application on which priority is claimed under 35 U.S.C. §§ 199, 365(a) or 365(b).) For a c-i-p application, applicant should review whether any claim in the patent that will issue is supported by an earlier application and, if not, the applicant should consider canceling the reference to the earlier filed application. The term of a patent is not based on a claim-by-claim approach. See Notice of April 14, 1995, 60 Fed. Reg. 20,195, at 20,205.

**WARNING:** When the last day of pendency of a provisional application falls on a Saturday, Sunday, or Federal holiday within the District of Columbia, any nonprovisional application claiming benefit of the provisional application **must** be filed prior to the Saturday, Sunday, or Federal holiday within the District of Columbia. See 37 C.F.R. § 1.78(a)(3).

- ☐ The new application being transmitted claims the benefit of prior U.S. application(s). Enclosed are ADDED PAGES FOR NEW APPLICATION TRANSMITTAL WHERE BENEFIT OF PRIOR U.S. APPLICATION(S) CLAIMED.

### 3. Papers Enclosed

**A.** Required for filing date under 37 C.F.R. § 1.53(b) (Regular) or 37 C.F.R. § 1.153 (Design) Application

18 Pages of specification

8 Pages of claims

12 Sheets of drawings

**WARNING:** **DO NOT** submit original drawings. A high quality copy of the drawings should be supplied when filing a patent application. The drawings that are submitted to the Office must be on strong, white, smooth, and non-shiny paper and meet the standards according to § 1.84. If corrections to the drawings are necessary, they should be made to the original drawing and a high-quality copy of the corrected original drawing then submitted to the Office. Only one copy is required or desired. For comments on proposed then-new 37 C.F.R. § 1.84, see Notice of March 9, 1988 (1990 O.G. 57-62).

**NOTE:** "Identifying indicia, if provided, should include the application number or the title of the invention, inventor's name, docket number (if any), and the name and telephone number of a person to call if the Office is unable to match the drawings to the proper application. This information should be placed on the back of each sheet of drawing a minimum distance of 1.5 cm (5/8 inch) down from the top of the page . . ." 37 C.F.R. § 1.84(c).

(complete the following, if applicable)

- ☐ The enclosed drawing(s) are photograph(s), and there is also attached a "PETITION TO ACCEPT PHOTOGRAPH(S) AS DRAWING(S)." 37 C.F.R. § 1.84(b).
- ☐ formal
- ☒ informal

### B. Other Papers Enclosed

     Pages of declaration and power of attorney

  1   Pages of abstract

  1   Other (Title Page)

### 4. Additional papers enclosed

- ☐ Amendment to claims
- ☐ Cancel in this application claims \_\_\_\_\_ before calculating the filing fee. (At least one original independent claim must be retained for filing purposes.)
- ☐ Add the claims shown on the attached amendment. (Claims added have been numbered consecutively following the highest numbered original claims.)
- ☐ Preliminary Amendment
- ☐ Information Disclosure Statement (37 C.F.R. § 1.98)
- ☐ Form PTO-1449 (PTO/SB/08A and 08B)
- ☐ Citations

- ☐ Declaration of Biological Deposit
- ☐ Submission of "Sequence Listing," computer readable copy and/or amendment pertaining thereto for biotechnology invention containing nucleotide and/or amino acid sequence.
- ☐ Authorization of Attorney(s) to Accept and Follow Instructions from Representative
- ☐ Special Comments
- ☐ Other

**5. Declaration or oath (including power of attorney)**

**NOTE:** *A newly executed declaration is not required in a continuation or divisional application provided that the prior nonprovisional application contained a declaration as required, the application being filed is by all or fewer than all the inventors named in the prior application, there is no new matter in the application being filed, and a copy of the executed declaration filed in the prior application (showing the signature or an indication thereon that it was signed) is submitted. The copy must be accompanied by a statement requesting deletion of the names of person(s) who are not inventors of the application being filed. If the declaration in the prior application was filed under § 1.47, then a copy of that declaration must be filed accompanied by a copy of the decision granting § 1.47 status or, if a nonsigning person under § 1.47 has subsequently joined in a prior application, then a copy of the subsequently executed declaration must be filed. See 37 C.F.R. §§ 1.63(d)(1)-(3).*

**NOTE:** *A declaration filed to complete an application must be executed, identify the specification to which it is directed, identify each inventor by full name including family name and at least one given name, without abbreviation together with any other given name or initial, and the residence, post office address and country or citizenship of each inventor, and state whether the inventor is a sole or joint inventor. 37 C.F.R. § 1.63(a)(1)-(4).*

**NOTE:** *"The inventorship of a nonprovisional application is that inventorship set forth in the oath or declaration as prescribed by § 1.62, except as provided for in § 1.53(d)(4) and § 1.63(d). If an oath or declaration as prescribed by § 1.63 is not filed during the pendency of a nonprovisional application, the inventorship is that inventorship set forth in the application papers filed pursuant to § 1.53(b), unless a petition under this paragraph accompanied by the fee set forth in § 1.17(f) is filed supplying or changing the name or names of the inventor or inventors." 37 C.F.R. § 1.41(a)(1).*

☐ Enclosed

Executed by

(check all applicable boxes)

- ☐ inventor(s).
- ☐ legal representative of inventor(s). 37 C.F.R. §§ 1.42 or 1.43.
- ☐ joint inventor or person showing a proprietary interest on behalf of inventor who refused to sign or cannot be reached.
- ☐ This is the petition required by 37 C.F.R. § 1.47 and the statement required by 37 C.F.R. § 1.47 is also attached. See item 13 below for fee.
- ☒ Not Enclosed

**NOTE:** *Where the filing is a completion in the U.S. of an International Application or where the completion of the U.S. application contains subject matter in addition to the International Application, the application may be treated as a continuation or continuation-in-part, as the case may be, utilizing ADDED PAGE FOR NEW APPLICATION TRANSMITTAL WHERE BENEFIT OF PRIOR U.S. APPLICATION CLAIMED.*

- ☐ Application is made by a person authorized under 37 C.F.R. § 1.41(c) on behalf of all the above named inventor(s).

(The declaration or oath, along with the surcharge required by 37 C.F.R. § 1.16(e) can be filed subsequently).

- ☐ Showing that the filing is authorized.  
(not required unless called into question. 37 C.F.R. § 1.41(d))

## 6. Inventorship Statement

**WARNING:** If the named inventors are each not the inventors of all the claims an explanation, including the ownership of the various claims at the time the last claimed invention was made, should be submitted.

The inventorship for all the claims in this application are:

- ☒ The same.

or

- ☐ Not the same. An explanation, including the ownership of the various claims at the time the last claimed invention was made,  
☐ is submitted.  
☐ will be submitted

## 7. Language

**NOTE:** An application including a signed oath or declaration may be filed in a language other than English. An English translation of the non-English language application and the processing fee of \$130.00 required by 37 C.F.R. § 1.17(k) is required to be filed with the application, or within such time as may be set by the Office. 37 C.F.R. § 1.52(d).

- ☒ English  
☐ Non English  
☐ The attached translation includes a statement that the translation is accurate.  
37 C.F.R. § 1.52(d).

## 8. Assignment

- ☒ An assignment of the invention to Nokia Corporation

- ☐ is attached. A separate ☐ "COVER SHEET FOR ASSIGNMENT (DOCUMENT) ACCOMPANYING NEW PATENT APPLICATION" or ☐ FORM PTO 1595 is also attached.  
☒ will follow.

**NOTE:** "If an assignment is submitted with a new application, send two separate letters-one for the application and one for the assignment." Notice of May 4, 1990 (1114 O.G. 77-78).

**WARNING:** A newly executed "CERTIFICATE UNDER 37 C.F.R. § 3.73(b)" must be filed when a continuation-in-part application is filed by an assignee. Notice of April 30, 1993, 1150 O.G. 62-64.

## 9. Certified Copy

Certified copy(ies) of application(s)

Country	Appln. No.	Filed
Country	Appln. No.	Filed
Country	Appln. No.	Filed

from which priority is claimed

- ☐ is (are) attached.  
☐ will follow.

NOTE: The foreign application forming the basis for the claim for priority must be referred to in the oath or declaration. 37 C.F.R. § 1.55(a) and 1.63.

NOTE: This item is for any foreign priority for which the application being filed directly relates. If any parent U.S. application or International Application from which this application claims benefit under 35 U.S.C. § 120 is itself entitled to priority from a prior foreign application, then complete item 18 on the ADDED PAGES FOR NEW APPLICATION TRANSMITTAL WHERE BENEFIT OF PRIOR U.S. APPLICATION(S) CLAIMED.

## 10. Fee Calculation (37 C.F.R. § 1.16)

- A. ☐ Regular application

CLAIMS AS FILED			
Number filed	Number Extra	Rate	Basic Fee 37 C.F.R. § 1.16(a) \$690.00
Total Claims (37 C.F.R. § 1.16(c))	-20 = 0	x \$18.00 =	
Independent Claims (37 C.F.R. § 1.16(b))	- 3 = 0	x \$78.00 =	
Multiple dependent claim(s), if any (37 C.F.R. § 1.16(d))		+ \$260.00	

- ☐ Amendment canceling extra claims is enclosed.  
☐ Amendment deleting multiple-dependencies is enclosed.  
☐ Fee for extra claims is not being paid at this time.

NOTE: If the fees for extra claims are not paid on filing, they must be paid or the claims canceled by amendment, prior to the expiration of the time period set for response by the Patent and Trademark Office in any notice of fee deficiency. 37 C.F.R. § 1.16(d).

Filing Fee Calculation \$ \_\_\_\_\_

- B. ☐ Design application  
(\$310.00 – 37 C.F.R. § 1.16(f))

Filing Fee Calculation \$ \_\_\_\_\_

C. ☐ Plant application

(\$480.00 - 37 C.F.R. § 1.16(g))

Filing Fee Calculation

\$ \_\_\_\_\_

**11. Small Entity Statement(s)**

- ☐ Statement(s) that this is a filing by a small entity under 37 C.F.R. §§ 1.9 and 1.27 is (are) attached.

**WARNING:** "Status as a small entity must be specifically established in each application or patent in which the status is available and desired. Status as a small entity in one application or patent does not affect any other application or patent, including applications or patents which are directly or indirectly dependent upon the application or patent in which the status has been established. The refiling of an application under § 1.53 as a continuation, division, or continuation-in-part (including a continued prosecution application under § 1.53(d)), or the filing of a reissue application requires a new determination as to continued entitlement to small entity status for the continuing or reissue application. A nonprovisional application claiming benefit under 35 U.S.C. § 119(e), 120, 121, or 365(c) of a prior application, or a reissue application may rely on a statement filed in the prior application or in the patent if the nonprovisional application or the reissue application includes a reference to the statement in the prior application or in the patent or includes a copy of the statement in the prior application or in the patent and status as a small entity is still proper and desired. The payment of the small entity basic statutory filing fee will be treated as such a reference for purposes of this section." 37 C.F.R. § 1.28(a)(2).

**WARNING:** "Small entity status must not be established when the person or persons signing the . . . statement can unequivocally make the required self-certification." M.P.E.P., § 509.03, 6<sup>th</sup> ed., rev. 2, July 1996 (emphasis added).

(complete the following, if applicable)

- ☐ Status as a small entity was claimed in prior application

\_\_\_\_\_, filed on \_\_\_\_\_, from which  
benefit is being claimed for this application under:

35 U.S.C. § ☐ 119(e),

☐ 120,

☐ 121,

☐ 365(c),

and which status as a small entity is still proper and desired.

- ☐ A copy of the statement in the prior application is included.

Filing Fee Calculation (50% of A, B, or C above)

\$ \_\_\_\_\_

**NOTE:** Any excess of the full fee paid will be refunded if a small entity statement and a refund request are filed within 2 months of the date of timely payment of a full fee. The two-month period is not extendable under § 1.136. 37 C.F.R. § 1.28(a).

**12. Request for International-Type Search (37 C.F.R. § 1.104(d))**

(complete, if applicable)

- ☐ Please prepare an international-type search report for this application at the time when national examination on the merits takes place.

**13. Fee Payment Being Made at This Time**

☒ Not Enclosed

☒ No filing fee is to be paid at this time.  
(This and the surcharge required by 37 C.F.R. § 1.16(e) can be paid subsequently.)

☐ Enclosed

☐ Filing fee \$ \_\_\_\_\_

☐ Recording assignment  
(\$40.00 – 37 C.F.R. § 1.21(h))  
(See attached "COVER SHEET  
FOR ASSIGNMENT ACCOMPANYING  
NEW APPLICATION.") \$ \_\_\_\_\_

☐ Petition fee for filing by other than all the  
inventors or person on behalf of the inventor  
where inventor refused to sign or cannot be  
reached.  
(\$130.00 – 37 C.F.R. §§ 1.47 and 1.17(i)) \$ \_\_\_\_\_

☐ For processing an application with a  
specification in a non-English language  
(\$130.00; 37 C.F.R. §§ 1.52(d) and 1.17(k)) \$ \_\_\_\_\_

☐ Processing and retention fee  
(\$130.00, 37 C.F.R. §§ 1.52(d) and 1.21(l)) \$ \_\_\_\_\_

☐ Fee for international-type search report  
(\$40.00; 37 C.F.R. § 1.21(e)) \$ \_\_\_\_\_

NOTE: 37 C.F.R. § 1.21(f) establishes a fee for processing and retaining any application that is abandoned for failing to complete the application pursuant to 37 C.F.R. § 1.53(f) and this, as well as the changes to 37 C.F.R. §§ 1.53 and 1.78(a)(1), indicate that in order to obtain the benefit of a prior U.S. application, either the basic filing fee must be paid, or the processing and retention fee of § 1.21(f) must be paid, within 1 year from the notification under § 53(f).

**Total fees enclosed** \$ \_\_\_\_\_

**14. Method of Payment of Fees**

☐ Check in the amount of \$ \_\_\_\_\_

☐ Charge Account No. \_\_\_\_\_ in the amount of \$ \_\_\_\_\_  
A duplicate of this transmittal is attached.

NOTE: Fees should be itemized in such a manner that it is clear for which purpose the fees are paid. 37 C.F.R. § 1.22(b).



## 15. Authorization to Charge Additional Fees

**WARNING:** If no fees are to be paid on filing, the following items should not be completed.

**WARNING:** Accurately count claims, especially multiple dependent claims, to avoid unexpected high charges, if extra claim charges are authorized.

☐ The Commissioner is hereby authorized to charge the following additional fees by this paper and during the entire pendency of this application to Account No. \_\_\_\_\_.

- ☐ 37 C.F.R. § 1.16(a), (f), or (g) (filing fees)
- ☐ 37 C.F.R. § 1.16(b), (c), and (d) (presentation of extra claims)

**NOTE:** Because additional fees for excess or multiple dependent claims not paid on filing or on later presentation must only be paid or these claims canceled by amendment prior to the expiration of the time period set for response by the P.T.O. in any notice of fee deficiency (37 C.F.R. § 1.16(d)), it might be best not to authorize the P.T.O. to charge additional claim fees, except possibly when dealing with amendments after final action.

- ☐ 37 C.F.R. § 1.16(e) (surcharge for filing the basic filing fee and/or declaration on a date later than the filing date of the application)
- ☐ 37 C.F.R. § 1.17(a)(1)-(5) (extension fees pursuant to § 1.136(a))
- ☐ 37 C.F.R. § 1.17 (application processing fees)

**WARNING:** "...A written request may be submitted in an application that is an authorization to treat any concurrent or future reply, requiring a petition for an extension of time under this paragraph for its timely submission, as incorporating a petition for extension of time for the appropriate length of time. An authorization to charge all required fees, fees under § 1.17, or all required extension of time fees will be treated as a constructive petition for an extension of time in any concurrent or future reply requiring a petition for an extension of time under this paragraph for its timely submission. Submission of the fee set forth in § 1.17(a) will also be treated as a constructive petition for an extension of time in any concurrent reply requiring a petition for an extension of time under this paragraph for its timely submission." 37 C.F.R. § 1.136(a)(3).

- ☐ 37 C.F.R. § 1.18 (issue fee at or before mailing of Notice of Allowance, pursuant to 37 C.F.R. § 1.311(b))

**NOTE:** Where an authorization to charge the issue fee to a deposit account has been filed before the mailing of a Notice of Allowance, the issue fee will be automatically charged to the deposit account at the time of mailing the Notice of Allowance. 37 C.F.R. § 1.311(b).

**NOTE:** 37 C.F.R. § 1.28(b) requires "Notification of any change in status resulting in loss of entitlement to small entity status must be filed in the application . . . prior to paying, or at the time of paying, . . . issue fee." From the wording of 37 C.F.R. § 1.28(b), (a) notification of change of status must be made even if the fee is paid as "other than a small entity" and (b) no notification is required if the change is to another small entity.

16. Instructions as to Overpayment

NOTE: "... Amounts of twenty-five dollars or less will not be returned unless specifically requested within a reasonable time, nor will the payer be notified of such amounts; amounts over twenty-five dollars may be returned by check or, if requested, by credit to a deposit account." 37 C.F.R. § 1.26(a).

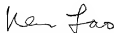
- ☐ Credit Account No. \_\_\_\_\_  
☐ Refund

Date: Oct. 31, 2000

Reg. No. 40,061

Tel. No. (203) 261-1234

Customer No. 004955



**SIGNATURE OF PRACTITIONER**

Kenneth Q. Lao

(type or print name of practitioner  
Ware, Fressola, Van Der Sluys &  
Adolphson LLP  
755 Main Street

P.O. (Correspondence) Address  
P.O. Box 224  
Monroe, CT 06468

☐ **Incorporation by reference of added pages**

(check the following item if the application in this transmittal claims the benefit of prior U.S. application(s) (including an international application entering the U.S. stage as a continuation, divisional or C-I-P application) and complete and attach the ADDED PAGES FOR NEW APPLICATION TRANSMITTAL WHERE BENEFIT OF PRIOR U.S. APPLICATION(S) CLAIMED.)

☐ Plus Added Pages for New Application Transmittal Where Benefit of Prior U.S. Application(s) Claimed

Number of pages added \_\_\_\_\_

☐ Plus Added Pages for Papers Referred to in Item 4 Above

Number of pages added \_\_\_\_\_

☐ Plus added pages deleting names of inventor(s) named in prior application(s) who is/are no longer inventor(s) of the subject matter claimed in this application.

Number of pages added \_\_\_\_\_

☐ Plus "Assignment Cover Letter Accompanying New Application"

Number of pages added \_\_\_\_\_

☒ **Statement Where No Further Pages Added**

*(if no further pages form a part of this Transmittal, then end this Transmittal with this page and check the following item.*

☒ This transmittal ends with this page.

PATENT  
Attorney Docket No. 944-003.040

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

PATENT APPLICATION

of

Jari MÄKINEN,  
Hannu MIKKOLA,  
Janne VAINIO  
and  
Jani ROTOLA-PUKKILA

for a

**METHOD AND SYSTEM FOR SPEECH FRAME  
ERROR CONCEALMENT IN SPEECH DECODING**

Express Mail Label #EL628639055US

0902540-103100

## METHOD AND SYSTEM FOR SPEECH FRAME ERROR CONCEALMENT IN SPEECH DECODING

### Field of the Invention

5 The present invention relates generally to the decoding of speech signals from an encoded bit stream and, more particularly, to the concealment of corrupted speech parameters when errors in speech frames are detected during speech decoding.

### Background of the Invention

10 Speech and audio coding algorithms have a wide variety of applications in communication, multimedia and storage systems. The development of the coding algorithms is driven by the need to save transmission and storage capacity while maintaining the high quality of the synthesized signal. The complexity of the coder is limited by, for example, the processing power of the application platform. In some applications, for example, voice storage, the encoder may be highly complex, while the decoder should be as simple as possible.

15 Modern speech codecs operate by processing the speech signal in short segments called frames. A typical frame length of a speech codec is 20 ms, which corresponds to 160 speech samples, assuming an 8 kHz sampling frequency. In the wide band codecs, the typical frame length of 20 ms corresponds to 320 speech samples, assuming a 16 kHz sampling frequency. The frame may be further divided into a number of sub-frames. For every frame, the encoder determines a parametric representation of the input signal. The parameters are quantized and transmitted through a communication channel (or stored in a storage medium) in a digital form. The decoder produces a synthesized speech signal based on the received parameters, as shown in Figure 1.

20 A typical set of extracted coding parameters includes spectral parameters (such as Linear Predictive Coding (LPC) parameters) to be used in short term prediction of the signal, parameters to be used for long term prediction (LTP) of the signal, various gain parameters, and excitation parameters. The LTP parameter is closely related to the fundamental frequency of the speech signal. This parameter is often known as a so-called pitch-lag

parameter, which describes the fundamental periodicity in terms of speech samples. Also, one of the gain parameters is very much related to the fundamental periodicity and so it is called LTP gain. The LTP gain is a very important parameter in making the speech as natural as possible. The description of the coding parameters above fits in general terms with a variety of speech codecs, including the so-called Code-Excited Linear Prediction (CELP) codecs, which have for some time been the most successful speech codecs.

Speech parameters are transmitted through a communication channel in a digital form. Sometimes the condition of the communication channel changes, and that might cause errors to the bit stream. This will cause frame errors (bad frames), i.e., some of the parameters describing a particular speech segment (typically 20 ms) are corrupted. There are two kinds of frame errors: totally corrupted frames and partially corrupted frames. These frames are sometimes not received in the decoder at all. In the packet-based transmission systems, like in normal internet connections, the situation can arise when the data packet will never reach the receiver, or the data packet arrives so late that it cannot be used because of the real time nature of spoken speech. The partially corrupted frame is a frame that does arrive to the receiver and can still contain some parameters that are not in error. This is usually the situation in a circuit switched connection like in the existing GSM connection. The bit-error rate (BER) in the partially corrupted frames is typically around 0.5-5%.

From the description above, it can be seen that the two cases of bad or corrupted frames will require different approaches in dealing with the degradation in reconstructed speech due to the loss of speech parameters.

The lost or erroneous speech frames are consequences of the bad condition of the communication channel, which causes errors to the bit stream. When an error is detected in the received speech frame, an error correction procedure is started. This error correction procedure usually includes a substitution procedure and muting procedure. In the prior art, the speech parameters of the bad frame are replaced by attenuated or modified values from the previous good frame. However, some parameters (such as excitation in CELP parameters) in the corrupted frame may still be used for decoding.

Figure 2 shows the principle of the prior-art method. As shown in Figure 2, a buffer labeled "parameter history" is used to store the speech parameters of the last good frame.

When a bad frame is detected, the Bad Frame Indicator (BFI) is set to 1 and the error concealment procedure is started. When the BFI is not set (BFI=0), the parameter history is updated and speech parameters are used for decoding without error concealment. In the prior-art system, the error concealment procedure uses the parameter history for concealing the lost or erroneous parameters in the corrupted frames. Some speech parameters may be used from the received frame even though it is classified as a bad frame (BFI=1). For example, in a GSM Adaptive Multi-Rate (AMR) speech codec (ETSI specification 06.91), the excitation vector from the channel is always used. When the speech frames are totally lost frames (e.g., in some IP-based transmission systems), no parameters will be used from the received bad frame. In some cases, no frame will be received, or the frame will arrive so late that it has to be classified as a lost frame.

In a prior-art system, LTP-lag concealment uses the last good LTP-lag value with a slightly modified fractional part, and spectral parameters are replaced by the last good parameters slightly shifted towards constant mean. The gains (LTP and fixed codebook) may usually be replaced by the attenuated last good value or by the median of several last good values. The same substituted speech parameters are used for all sub-frames with slight modification to some of them.

The prior-art LTP concealment may be adequate for stationary speech signals, for example, voiced or stationary speech. However, for non-stationary speech signals, the prior-art method may cause unpleasant and audible artifacts. For example, when the speech signal is unvoiced or non-stationary, simply substituting the lag value in the bad frame with the last good lag value has the effect of generating a short voiced-speech segment in the middle of an unvoiced-speech burst (See Figure 10). The effect, as known as the Abing≡ artifact, can be annoying.

It is advantageous and desirable to provide a method and system for error concealment in speech decoding to improve the speech quality.

#### Summary of the Invention

The present invention takes advantage of the fact that there is a recognizable relationship among the long-term prediction (LTP) parameters in the speech signals. In

particular, the LTP-lag has a strong correlation with the LTP-gain. When the LTP-gain is high and reasonably stable, the LTP-lag is typically very stable and the variation between adjacent lag values is small. In that case, the speech parameters are indicative of a voiced-speech sequence. When the LTP-gain is low or unstable, the LTP-lag is typically unvoiced,  
5 and the speech parameters are indicative of an unvoiced-speech sequence. Once the speech sequence is classified as stationary (voiced) or non-stationary (unvoiced), the corrupted or bad frame in the sequence can be processed differently.

Accordingly, the first aspect of the present invention is a method for concealing errors in an encoded bit stream indicative of speech signals received in a speech decoder, wherein the encoded bit stream includes a plurality of speech frames arranged in speech sequences, and the speech frames include at least one corrupted frame preceded by one or more non-corrupted frames, wherein the corrupted frame includes a first long-term prediction lag value and a first long-term prediction gain value, and the non-corrupted frames include second long-term prediction lag values and second long-term prediction gain values, and wherein the  
15 second long-term prediction lag values include a last long-term prediction lag value, and the second long-term prediction gain values include a last long-term prediction gain value, and the speech sequences include stationary and non-stationary speech sequences, and wherein the corrupted frame can be partially corrupted or totally corrupted. The method comprises the steps of:

- 20 determining whether the first long-term prediction lag value is within or outside an upper limit and a lower limit determined based on the second long-term prediction lag values;  
replacing the first long-term prediction lag value in the partially corrupted frame with a third lag value, when the first long-term prediction lag value is outside the upper and lower limits; and
- 25 retaining the first long-term prediction lag value in the partially corrupted frame when the first long-term prediction lag value is within the upper and lower limits.

Alternatively, the method comprises the steps of:

- determining whether the speech sequence in which the corrupted frame is arranged is stationary or non-stationary, based on the second long-term prediction gain values;
- 30 when the speech sequence is stationary, replacing the first long-term prediction lag



value in the corrupted frame with the last long-term prediction lag value; and

when the speech sequence is non-stationary, replacing the first long-term prediction lag value in the corrupted frame with a third long-term prediction lag value determined based on the second long-term prediction lag values and an adaptively-limited random lag jitter, and  
5 replacing the first long-term prediction gain value in the corrupted frame with a third long-term prediction gain value determined based on the second long-term prediction gain values and an adaptively-limited random gain jitter.

Preferably, the third long-term prediction lag value is calculated based at least partially on a weighted median of the second long-term prediction lag values, and the adaptively-limited random lag jitter is a value bound by limits determined based on the second long-term prediction lag values.

Preferably, the third long-term prediction gain value is calculated based at least partially on a weighted median of the second long-term prediction gain values, and the adaptively-limited random gain jitter is a value bound by limits determined based on the second long-term prediction gain values.

Alternatively, the method comprises the steps of:

determining whether the corrupted frame is partially corrupted or totally corrupted;

replacing the first long-term prediction lag value in the corrupted frame with a third lag value if the corrupted frame is totally corrupted, wherein when the speech sequence in which the totally corrupted frame is arranged is stationary, set the third lag value equal to the last long-term prediction lag value, and when said speech sequence is non-stationary, determining the third lag value based on the second long-term prediction values and an adaptively-limited random lag jitter; and

replacing the first long-term prediction lag value in the corrupted frame with a fourth lag value if the corrupted frame is partially corrupted., wherein when the speech sequence in which the partially corrupted frame is arranged in stationary, set the fourth lag value equal to the last long-term prediction lag value, and when said speech sequence is non-stationary set the fourth lag value based on a decoded long-term prediction lag value searched from an adaptive codebook associated with the non-corrupted frame preceding the corrupted frame, when said speech sequence is non-stationary.

The second aspect of the present invention is a speech signal transmitter and receiver system for encoding speech signals in an encoded bit stream and decoding the encoded bit stream into synthesized speech, wherein the encoded bit stream includes a plurality of speech frames arranged in speech sequences, and the speech frames include at least one corrupted frame preceded by one or more non-corrupted frames, wherein the corrupted frame is indicated by a first signal and includes a first long-term prediction lag value and a first long-term prediction gain value, and the non-corrupted frames include second long-term prediction lag values and second long-term prediction gain values, and wherein the second long-term prediction lag values include a last long-term prediction lag value, and the second long-term prediction gain values include a last long-term prediction gain value, and the speech sequences include stationary and non-stationary speech sequences. The system comprises:

a first mechanism, responsive to the first signal, for determining whether the speech sequence in which the corrupted frame is arranged is stationary or non-stationary, based on the second long-term prediction gain values, and for providing a second signal indicative of whether the speech sequence is stationary or non-stationary; and

a second mechanism, responsive to the second signal, for replacing the first long-term prediction lag value in the corrupted frame with the last long-term prediction lag value when the speech sequence is stationary, and replacing the first long-term prediction lag value and the first long-term gain value in the corrupted frame with a third long-term prediction lag value and a third long-term prediction gain value, respectively, when the speech sequence is non-stationary, wherein the third long-term prediction lag value is determined based on the second long-term prediction lag values and an adaptively-limited random lag jitter, and the third long-term prediction gain value is determined based on the second long-term prediction gain values and an adaptively-limited random gain jitter.

Preferably, the third long-term prediction lag value is calculated based at least partially on a weighted median of the second long-term prediction lag values, and the adaptively-limited random lag jitter is a value bound by limits determined based on the second long-term prediction lag values.

Preferably, the third long-term prediction gain value is calculated based at least partially on a weighted median of the second long-term prediction gain values, and the

adaptively-limited random gain jitter is a value bound by limits determined based on the second long-term prediction gain values.

The third aspect of the present invention is a decoder for synthesizing speech from an encoded bit stream, wherein the encoded bit stream includes a plurality of speech frames arranged in speech sequences, and the speech frames include at least one corrupted frame preceded by one or more non-corrupted frames, wherein the corrupted frame is indicated by a first signal and includes a first long-term prediction lag value and a first long-term prediction gain value, and the non-corrupted frames include second long-term prediction lag values and second long-term prediction gain values, and wherein the second long-term prediction lag values include a last long-term prediction lag value and the second long-term prediction gain values include a last long-term prediction gain value and the speech sequences include stationary and non-stationary speech sequences. The decoder comprises:

a first mechanism, responsive to the first signal, for determining whether the speech sequence in which the corrupted frame is arranged is stationary or non-stationary, based on the second long-term prediction gain values, and for providing a second signal indicative of whether the speech sequence is stationary or non-stationary; and

a second mechanism, responsive to the second signal, for replacing the first long-term prediction lag value in the corrupted frame with the last long-term prediction lag value when the speech sequence is stationary, and replacing the first long-term prediction lag value and the first long-term gain value in the corrupted frame with a third long-term prediction lag value and a third long-term prediction gain value, respectively, when the speech sequence is non-stationary, wherein the third long-term prediction lag value is determined based on the second long-term prediction lag values and an adaptively-limited random lag jitter, and the third long-term prediction gain value is determined based on the second long-term prediction gain values and an adaptively-limited random gain jitter.

The fourth aspect of the present invention is a mobile station, which is arranged to receive an encoded bit stream containing speech data indicative of speech signals, wherein the encoded bit stream includes a plurality of speech frames arranged in speech sequences, and the speech frames include at least one corrupted frame preceded by one or more non-corrupted frames, wherein the corrupted frame is indicated by a first signal and includes a

first long-term prediction lag value and a first long-term prediction gain value, and the non-corrupted frames include second long-term prediction lag values and second long-term prediction gain values, and wherein the second long-term prediction lag values include a last long-term prediction lag value and the second long-term prediction gain values include a last long-term prediction gain value and the speech sequences include stationary and non-stationary speech sequences. The mobile station comprises:

a first mechanism, responsive to the first signal, for determining whether the speech sequence in which the corrupted frame is arranged is stationary or non-stationary, based on the second long-term prediction gain values, and for providing a second signal indicative of whether the speech sequence is stationary or non-stationary; and

a second mechanism, responsive to the second signal, for replacing the first long-term prediction lag value in the corrupted frame with the last long-term prediction lag value when the speech sequence is stationary, and replacing the first long-term prediction lag value and the first long-term gain value in the corrupted frame with a third long-term prediction lag value and a third long-term prediction gain value, respectively, when the speech sequence is non-stationary, wherein the third long-term prediction lag value is determined based on the second long-term prediction lag values and an adaptively-limited random lag jitter, and the third long-term prediction gain value is determined based on the second long-term prediction gain values and an adaptively-limited random gain jitter.

The fifth aspect of the present invention is an element in a telecommunication network, which is arranged to receive an encoded bit stream containing speech data from a mobile station, wherein the speech data includes a plurality of speech frames arranged in speech sequences, and the speech frames include at least one corrupted frame preceded by one or more non-corrupted frames, wherein the corrupted frame is indicated by a first signal and includes a first long-term prediction lag value and a first long-term prediction gain value, and the non-corrupted frames include second long-term prediction lag values and second long-term prediction gain values, and wherein the second long-term prediction lag values include a last long-term prediction lag value and the second long-term prediction gain values include a last long-term prediction gain value and the speech sequences include stationary and non-stationary speech sequences. The element comprises:

a first mechanism, responsive to the first signal, for determining whether the speech sequence in which the corrupted frame is arranged is stationary or non-stationary, based on the second long-term prediction gain values, and for providing a second signal indicative of whether the speech sequence is stationary or non-stationary; and

a second mechanism, responsive to the second signal, for replacing the first long-term prediction lag value in the corrupted frame with the last long-term prediction lag value when the speech sequence is stationary, and replacing the first long-term prediction lag value and the first long-term gain value in the corrupted frame with a third long-term prediction lag value and a third long-term prediction gain value, respectively, when the speech sequence is non-stationary, wherein the third long-term prediction lag value is determined based on the second long-term prediction lag values and an adaptively-limited random lag jitter, and the third long-term prediction gain value is determined based on the second long-term prediction gain values and an adaptively-limited random gain jitter.

The present invention will become apparent upon reading the description taken in conjunction with Figures 3 to 11c.

#### Brief Description of the Drawings

Figure 1 is a block diagram illustrating a generic distributed speech codec, wherein the encoded bit stream containing speech data is conveyed from an encoder to a decoder via a communication channel or a storage medium.

Figure 2 is a block diagram illustrating a prior-art error concealment apparatus in a receiver.

Figure 3 is a block diagram illustrating the error concealment apparatus in a receiver, according to the present invention.

Figure 4 is a flow chart illustrating the method of error concealment according to the present invention.

Figure 5 is a diagrammatic representation of a mobile station, which includes an error concealment module, according to the present invention.

Figure 6 is a diagrammatic representation of a telecommunication network using a decoder, according to the present invention.

Figure 7 is a plot of LTP-parameters illustrating the lag and gain profiles in a voiced speech sequence.

Figure 8 is a plot of LTP-parameters illustrating the lag and gain profiles in an unvoiced speech sequence.

Figure 9 is a plot of LTP-lag values in a series of sub-frames illustrating the difference between the prior-art error concealment approach and the approach according to the present invention.

Figure 10 is another plot of LTP-lag values in a series of sub-frames illustrating the difference between the prior-art error concealment approach and the approach according to the present invention.

Figure 11a is a plot of speech signals illustrating an error-free speech sequence having the location of the bad frame of the speech channel, as shown in Figures 11b and 11c.

Figure 11b is a plot of speech signals illustrating the concealment of parameters in a bad frame according to the prior art approach.

Figure 11c is a plot of speech signals illustrating the concealment of parameters in a bad frame according to the present invention.

#### Best Mode for Carrying Out the Invention

Figure 3 illustrates a decoder **10**, which includes a decoding module **20** and an error concealment module **30**. The decoding module **20** receives a signal **140**, which is normally indicative of speech parameters **102** for speech synthesis. The decoding module **20** is known in the art. The error concealment module **30** is arranged to receive an encoded bit stream **100**, which includes a plurality of speech streams arranged in speech sequences. A bad-frame detection device **32** is used to detect corrupted frames in the speech sequences and provide a Bad-Frame-Indicator (BFI) signal **110** representing a BFI flag when a corrupted frame is detected. BFI is also known in the art. The BFI signal **110** is used to control two switches **40** and **42**. Normally, the speech frames are not corrupted and the BFI flag is 0. The terminal **S** is operatively connected to the terminal **0** in the switches **40** and **42**. The speech parameters **102** are conveyed to a buffer, or "parameter history" storage, **50** and the decoding module **20**

for speech synthesis. When a bad frame is detected by the bad-frame detection device 32, the BFI flag is set to 1. The terminal S is connected to the terminal 1 in the switches 40 and 42. Accordingly, the speech parameters 102 are provided to an analyzer 70, and the speech parameters needed for speech synthesis are provided by a parameter concealment module 60 to the decoding module 20. The speech parameters 102 typically include LPC parameters for short term prediction, excitation parameters, a long-term prediction (LTP) lag parameter, an LTP gain parameter and other gain parameters. The parameter history storage 50 is used to store the LTP-lag and LTP-gain of a number of non-corrupted speech frames. The contents of the parameter history storage 50 are constantly updated so that the last LTP-gain parameter and the last LTP-lag parameter stored in the storage 50 are those of the last non-corrupted speech frame. When a corrupted frame in a speech sequence is received in the decoder 10, the BFI flag is set to 1 and the speech parameters 102 of the corrupted frame are conveyed to the analyzer 70 through the switch 40. By comparing the LTP-gain parameter in the corrupted frame and the LTP-gain parameters stored in the storage 50, it is possible for the analyzer 70 to determine whether the speech sequence is stationary or non-stationary, based on the magnitude and its variation in the LTP-gain parameters in neighboring frames. Typically, in a stationary sequence, the LTP-gain parameters are high and reasonably stable, the LTP-lag value is stable and the variation in adjacent LTP-lag values is small, as shown in Figure 7. In contrast, in a non-stationary sequence, the LTP-gain parameters are low and unstable, and the LTP-lag is also unstable, as shown in Figure 8. The LTP-lag values are changing more or less randomly. Figure 7 shows the speech sequence for the word "viiniä". Figure 8 shows the speech sequence for the word "exhibition".

If the speech sequence that includes the corrupted frame is voiced or stationary, the last good LTP-lag is retrieved from the storage 50 and conveyed to the parameter concealment module 60. The retrieved good LTP-lag is used to replace the LTP-lag of the corrupted frame. Because the LTP-lag in a stationary speech sequence is stable and its variation is small, it is reasonable to use a previous LTP-lag with small modification to conceal the corresponding parameter in corrupted frame. Subsequently, an RX signal 104 causes the replacement parameters, as denoted by reference numeral 134, to be conveyed to

the decoding module 20 through the switch 42.

If the speech sequence that includes the corrupted frame is unvoiced or non-stationary, the analyzer 70 calculates a replacement LTP-lag value and a replacement LTP-gain value for parameter concealment. Because LTP-lag in a non-stationary speech sequence is unstable and its variation in adjacent frames is typically very large, parameter concealment should allow the LTP-lag in an error-concealed non-stationary sequence to fluctuate in a random fashion. If the parameters in the corrupted frame are totally corrupted, such as in a lost frame, the replacement LTP-lag is calculated by using a weighted median of the previous good LTP-lag values along with an adaptively-limited random jitter. The adaptively-limited random jitter is allowed to vary within limits calculated from the history of the LTP values, so that the parameter fluctuation in an error-concealed segment is similar to the previous good section of the same speech sequence.

An exemplary rule for LTP-lag concealment is governed by a set of conditions as follows:

If

$\text{minGain} > 0.5 \text{ AND } \text{LagDif} < 10; \text{ OR}$   
 $\text{lastGain} > 0.5 \text{ AND } \text{secondLastGain} > 0.5,$

then the last received good LTP-lag is used for the totally corrupted frame. Otherwise, *Update\_lag*, a weighted average of the LTP-lag buffer with randomization, is used for the totally corrupted frame. *Update\_lag* is calculated in a manner as described below:

The LTP-lag buffer is sorted and the three biggest buffer values are retrieved. The average of these three biggest values is referred to as the weighted average lag (*WAL*), and the difference from these biggest values is referred to as the weighted lag difference (*WLD*).

Let *RAND* be the randomization with the scale of  $(-WLD/2, WLD/2)$ , then

$\text{Update\_lag} = \text{WAL} + \text{RAND} (-WLD/2, WLD/2),$

wherein

*minGain* is the smallest value of the LTP-gain buffer;



*LagDif* is the difference between the smallest and the largest LTP-lag values;  
*lastGain* is the last received good LTP-gain; and  
*secondLastGain* is the second last received good LTP-gain.

If the parameters in the corrupted frame are partially corrupted, then the LTP-lag value in the corrupted frame is replaced accordingly. That the frame is partially corrupted is determined by a set of exemplary LTP-feature criteria given below:

If

- (1)  $LagDif < 10$  AND  $(minLag - 5) < T_{bf} < (maxLag + 5)$ ; OR  
(2)  $lastGain > 0.5$  AND  $secondLastGain > 0.5$  AND  $(lastLag - 10) < T_{bf} < (lastLag + 10)$ ;

OR

- (3)  $minGain < 0.4$  AND  $lastGain = minGain$  AND  $minLag < T_{bf} < maxLag$ ; OR  
(4)  $LagDif < 70$  AND  $minLag < T_{bf} < maxLag$ ; OR  
(5)  $meanLag < T_{bf} < maxLag$

is true, then  $T_{bf}$  is used to replace the LTP-lag in the corrupted frame. Otherwise, the corrupted frame is treated as a totally corrupted frame, as described above. In the above conditions:

*maxLag* is the largest value of the LTP-lag buffer;

*meanLag* is the average of the LTP-lag buffer;

*minLag* is the smallest value of the LTP-lag buffer;

*lastLag* is the last received good LTP-lag value; and

$T_{bf}$  is a decoded LTP lag which is searched, when the BFI is set, from the adaptive codebook as if the BFI is not set.

Two examples of parameter concealment are shown in Figures 9 and 10. As shown, the profile of the replacement LTP-lag values in the bad frame, according to the prior art, is rather flat, but the profile of the replacement, according to the present invention, allows some fluctuation, similar to the error-free profile. The difference between the prior art approach

and the present invention is further illustrated in Figures 11b and 11c, respectively, based on the speech signals in an error-free channel, as shown in Figure 11a.

When the parameters in the corrupted frame are partially corrupted, the parameter concealment can be further optimized. In partially corrupted frames, the LTP-lags in the corrupted frames may still yield an acceptable synthesized speech segment. Accordingly to the GSM specifications, the BFI flag is set by a Cyclic Redundancy Check (CRC) mechanism or other error detection mechanisms. These error detection mechanisms detect errors in the most significant bits in the channel decoding process. Accordingly, even when only a few bits are erroneous, the error can be detected and the BFI flag is set accordingly. In the prior-art parameter concealment approach, the entire frame is discarded. As a result, information contained in the correct bits is thrown away.

Typically, in the channel decoding process, the BER per frame is a good indicator for the channel condition. When the channel condition is good, the BER per frame is small and a high percentage of the LTP-lag values in the erroneous frames are correct. For example, when the frame error rate (FER) is 0.2%, over 70% of the LTP-lag values are correct. Even when the FER reaches 3%, about 60% of the LTP-lag values are still correct. The CRC can accurately detect a bad frame and set the BFI flag accordingly. However, the CRC does not provide an estimation of the BER in the frame. If the BFI flag is used as the only criterion for parameter concealment, then a high percentage of the correct LTP-lag values could be wasted.

In order to prevent a large amount of correct LTP-lags from being thrown away, it is possible to adapt a decision criterion for parameter concealment based on the LTP history. It is also possible to use the FER, for example, as the decision criterion. If the LTP-lag meets the decision criterion, no parameter concealment is necessary. In that case, the analyzer 70 conveys the speech parameters 102, as received through the switch 40, to the parameter concealment module 60 which then conveys the same to the decoding module 20 through the switch 42. If the LTP-lag does not meet that decision criterion, then the corrupted frame is further examined using the LTP-feature criteria, as described hereinabove, for parameter concealment.

In stationary speech sequences, the LTP-lag is very stable. Whether most of the LTP-lag values in a corrupted frame are correct or erroneous can be correctly predicted with high

probability. Thus, it is possible to adapt a very strict criterion for parameter concealment. In non-stationary speech sequences, it may be difficult to predict whether the LTP-lag value in a corrupted frame is correct, because of the unstable nature of the LTP parameters. However, that the prediction is correct or wrong is less important in non-stationary speech than in stationary speech. While allowing erroneous LTP-lag values to be used in decoding stationary speech may cause the synthesized speech to be unrecognizable, allowing erroneous LTP-lag values to be used in decoding non-stationary speech usually only increases the audible artifacts. Thus, the decision criterion for parameter concealment in non-stationary speech can be relatively lax.

As mentioned earlier, the LTP-gain fluctuates greatly in non-stationary speech. If the same LTP-gain value from the last good frame is used repeatedly to replace the LTP-gain value of one or more corrupted frames in a speech sequence, the LTP-gain profile in the gain concealed segment will be flat (similar to the prior-art LTP-lag replacement, as shown in Figures 7 and 8), in stark contrast to the fluctuating profile of the non-corrupted frames. The sudden change in the LTP-gain profile may cause unpleasant audible artifacts. In order to minimize these audible artifacts, it is possible to allow the replacement LTP-gain value to fluctuate in the error-concealed segment. For this purpose, the analyzer 70 can be also used to determine the limits between which the replacement LTP-gain value is allowed to fluctuate based on the gain values in the LTP history.

LTP-gain concealment can be carried out in a manner as described below. When the BFI is set, a replacement LTP-gain value is calculated according to a set of LTP-gain concealment rules. The replacement LTP-gain is denoted as *Updated\_gain*.

- (1) If  $gainDif > 0.5$  AND  $lastGain = maxGain > 0.9$  AND  $subBF=1$ , then  
 $Updated\_gain = (secondLastGain + thirdLastGain)/2$ ;
- (2) If  $gainDif > 0.5$  AND  $lastGain = maxGain > 0.9$  AND  $subBF=2$ , then  
 $Updated\_gain = meanGain + randVar * (maxGain - meanGain)$ ;
- (3) If  $gainDif > 0.5$  AND  $lastGain = maxGain > 0.9$  AND  $subBF=3$ , then  
 $Updated\_gain = meanGain - randVar * (meanGain - minGain)$ ;
- (4) If  $gainDif > 0.5$  AND  $lastGain = maxGain > 0.9$  AND  $subBF=4$ , then

$$Updated\_gain = meanGain + randVar * (maxGain - meanGain);$$

In the previous conditions, *Updated\_gain* cannot be larger than *lastGain*. If the previous conditions cannot be met, the following conditions are used:

- (5) If *gainDif* > 0.5, then  
*Updated\_gain* = *lastGain*;
- (6) If *gainDif* < 0.5 AND *lastGain* = *maxGain*, then  
*Updated\_gain* = *meanGain*;
- (7) If *gainDIF* < 0.5, then  
*Updated\_gain* = *lastGain*,

Wherein

*meanGain* is the average of the LTP-gain buffer;  
*maxGain* is the largest value of the LTP-gain buffer;  
*minGain* is the smallest value of the LTP-gain buffer;  
*randVar* is a random value between 0 and 1,  
*gainDIF* is the difference between the smallest and the largest LTP-gain values in the LTP-gain buffer;  
*lastGain* is the last received good LTP-gain;  
*seconLastGain* is the second last received good LTP-gain;  
*thirdLastGain* is the third last received good LTP-gain; and  
*subBF* is the order of the subframe.

Figure 4 illustrates the method of error-concealment, according to the present invention. As the encoded bit stream is received at step 160, the frame is checked to see if it is corrupted at step 162. If the frame is not corrupted, then the parameter history of the speech sequence is updated at step 164, and the speech parameters of the current frame are decoded at step 166. The procedure then goes back to step 162. If the frame is bad or corrupted, the parameters are retrieved from the parameter history storage at step 170. Whether the corrupted frame is part of the stationary speech sequence or non-stationary

speech sequence is determined at step 172. If the speech sequence is stationary, the LTP-lag of the last good frame is used to replace the LTP-lag in the corrupted frame at step 174. If the speech sequence is non-stationary, a new lag value and new gain value are calculated based on the LTP history at step 180, and they are used to replace the corresponding parameters in the corrupted frame at step 182.

Figure 5 shows a block diagram of a mobile station 200 according to one exemplary embodiment of the invention. The mobile station comprises parts typical of the device, such as a microphone 201, keypad 207, display 206, earphone 214, transmit/receive switch 208, antenna 209 and control unit 205. In addition, the figure shows transmitter and receiver blocks 204, 211 typical of a mobile station. The transmitter block 204 comprises a coder 221 for coding the speech signal. The transmitter block 204 also comprises operations required for channel coding, deciphering and modulation as well as RF functions, which have not been drawn in Figure 5 for clarity. The receiver block 211 also comprises a decoding block 220 according to the invention. Decoding block 220 comprises an error concealment module 222 like the parameter concealment module 30 shown in Figure 3. The signal coming from the microphone 201, amplified at the amplification stage 202 and digitized in the A/D converter, is taken to the transmitter block 204, typically to the speech coding device comprised by the transmit block. The transmission signal, which is processed, modulated and amplified by the transmit block, is taken via the transmit/receive switch 208 to the antenna 209. The signal to be received is taken from the antenna via the transmit/receive switch 208 to the receiver block 211, which demodulates the received signal and decodes the deciphering and the channel coding. The resulting speech signal is taken via the D/A converter 212 to an amplifier 213 and further to an earphone 214. The control unit 205 controls the operation of the mobile station 200, reads the control commands given by the user from the keypad 207 and gives messages to the user by means of the display 206.

The parameter concealment module 30, according to the invention, can also be used in a telecommunication network 300, such as an ordinary telephone network, or a mobile station network, such as the GSM network. Figure 6 shows an example of a block diagram of such a telecommunication network. For example, the telecommunication network 300 can comprise

telephone exchanges or corresponding switching systems **360**, to which ordinary telephones **370**, base stations **340**, base station controllers **350** and other central devices **355** of telecommunication networks are coupled. Mobile stations **330** can establish connection to the telecommunication network via the base stations **340**. A decoding block **320**, which includes an error concealment module **322** similar to the error concealment module **30** shown in Figure 3, can be particularly advantageously placed in the base station **340**, for example. However, the decoding block **320** can also be placed in the base station controller **350** or other central or switching device **355**, for example. If the mobile station system uses separate transcoders, for example, between the base stations and the base station controllers, for transforming the coded signal taken over the radio channel into a typical 64 kbit/s signal transferred in a telecommunication system and vice versa, the decoding block **320** can also be placed in such a transcoder. In general, the decoding block **320**, including the parameter concealment module **322**, can be placed in any element of the telecommunication network **300**, which transforms the coded data stream into an uncoded data stream. The decoding block **320** decodes and filters the coded speech signal coming from the mobile station **330**, whereafter the speech signal can be transferred in the usual manner as uncompressed forward in the telecommunication network **300**.

It should be noted that the error concealment method of the present invention has been described with respect to stationary and non-stationary speech sequences, and that stationary speech sequences are usually voiced and non-stationary speech sequences are usually unvoiced. Thus, it will be understood that the disclosed method is applicable to error concealment in voiced and unvoiced speech sequences.

The present invention is applicable to CELP type speech codecs and can be adapted to other types of speech codecs as well. Thus, although the invention has been described with respect to a preferred embodiment thereof, it will be understood by those skilled in the art that the foregoing and various other changes, omissions and deviations in the form and detail thereof may be made without departing from the spirit and scope of this invention.

What is claimed is:

1. A method for concealing errors in an encoded bit stream indicative of speech signals received in a speech decoder, wherein the encoded bit stream includes a plurality of speech frames arranged in speech sequences, and the speech frames include at least one partially corrupted frame preceded by one or more non-corrupted frames, wherein the partially corrupted frame includes a first long-term prediction lag value and a first long-term prediction gain value, and the non-corrupted frames include second long-term prediction lag values and second long-term prediction gain values, and wherein the second long-term prediction lag values include a last long-term prediction lag value, and the second long-term prediction gain values include a last long-term prediction gain value, said method comprising the steps of:

providing an upper limit and a lower limit based on the second long-term prediction lag values;

determining whether the first long-term prediction lag value is within or outside the upper and lower limits;

replacing the first long-term prediction lag value in the partially corrupted frame with a third lag value, when the first long-term prediction lag value is outside the upper and lower limits; and

retaining the first long-term prediction lag value in the partially corrupted frame when the first long-term prediction lag value is within the upper and lower limits.

2. The method of claim 1, further comprising the step of replacing the first long-term prediction gain value in the partially corrupted frame with a third gain value, when the first long-term lag value is outside the upper and lower limits.

3. The method of claim 1, wherein the third lag value is calculated based the second long-term prediction lag values and an adaptively-limited random lag jitter bound by further limits determined based on the second long-term prediction lag values.

4. The method of claim 2, wherein the third gain value is calculated based on of the second long-term prediction gain values and an adaptively-limited random gain jitter bound

by limits determined based on the second long-term prediction gain values.

5. A method for concealing errors in an encoded bit stream indicative of speech signals received in a speech decoder, wherein the encoded bit stream includes a plurality of speech frames arranged in speech sequences, and the speech frames include at least one corrupted frame preceded by one or more non-corrupted frames, wherein the corrupted frame includes a first long-term prediction lag value and a first long-term prediction gain value, and the non-corrupted frames include second long-term prediction lag values and second long-term prediction gain values, and wherein the second long-term prediction lag values include a last long-term prediction lag value, and the second long-term prediction gain values include a last long-term prediction gain value and the speech sequences include stationary and non-stationary speech sequences, and wherein the corrupted frame can be a totally corrupted frame or a partially corrupted frame, said method comprising the steps of:

determining whether the corrupted frame is partially corrupted or totally corrupted;  
replacing the first long-term prediction lag value in the corrupted frame with a third lag value if the corrupted frame is totally corrupted; and  
replacing the first long-term prediction lag value in the corrupted frame with a fourth lag value if the corrupted frame is partially corrupted.

6. The method of claim 5, further comprising the steps of:

determining whether the speech sequence in which the partially corrupted frame is arranged is stationary or non-stationary;  
setting the fourth lag value equal to the last long-term prediction lag value, when said speech sequence is stationary; and  
determining the fourth lag value based on a decoded long-term prediction lag value searched from an adaptive codebook associated with the non-corrupted frame preceding the corrupted frame, when said speech sequence is non-stationary.

7. The method of claim 5, further comprising the steps of:

determining whether the speech sequence in which the totally corrupted frame is



arranged is stationary or non-stationary;

setting the third lag value equal to the last long-term prediction lag value, when said speech sequence is stationary; and

determining the third lag value based on the second long-term prediction values and an adaptively-limited random lag jitter, when said speech sequence is non-stationary.

8. The method of claim 5, wherein the second long-term prediction lag values further include a second last long-term prediction lag value and a third last long-term prediction lag value, and the second long-term prediction gain values further include a second last long-term prediction gain value and a third last long-term prediction gain value, said method further comprising the steps of:

determining minLag, which is the smallest lag value among the second long-term prediction lag values;

determining maxLag, which is the largest lag value among the second long-term prediction lag values;

determining meanLag, which is an average of the second long-term prediction lag values;

determining difLag, which is the difference of maxLad and minLag;

determining minGain, which is the smallest gain value among the second long-term prediction gain values;

determining maxGain, which is the largest gain value among the second long-term prediction gain values; and

determining meanGain, which is an average of the second long term gain values; wherein

if  $\text{difLag} < 10$ , and  $\text{minLag} < \text{the fourth lag value} < \text{maxLag} + 5$ ; or

if the last long-term prediction gain value is larger than 0.5, and the second last long-term prediction gain value is larger than 0.5, and the fourth lag value is smaller than a sum of the last long-term prediction value and 10, and a sum of the fourth lag value and 10 is larger than the last long-term prediction value; or

if  $\text{minGain} < 0.4$ , and the last long-term prediction gain value is equal to minGain, and

the fourth lag value is larger than minLag but smaller than maxLag; or

if  $\text{diffLag} < 70$ , and the fourth lag value is larger than minLag but smaller than maxLag; or

if the fourth lag value is larger than meanLag but smaller than maxLag; then the corrupted frame is determined as partially corrupted.

9. The method of claim 6, wherein when said speech sequence is non-stationary, said method further comprising the step of determining a frame-error rate of the speech frames such that

if the frame-error rate reaches a determined value, the fourth lag value is determined based on said decoded long-term prediction lag value, and

if the frame-error rate is smaller than the determined value, the fourth lag value is set equal to the last long-term prediction lag value.

10. The method of claim 5, wherein the stationary speech sequences include voiced sequences, and the non-stationary speech sequences include unvoiced sequences.

11. A speech signal transmitter and receiver system for encoding speech signals in an encoded bit stream and decoding the encoded bit stream into synthesized speech, wherein the encoded bit stream includes a plurality of speech frames arranged in speech sequences, and the speech frames include at least one corrupted frame preceded by one or more non-corrupted frames, wherein the corrupted frame includes frame a first long-term prediction lag value and a first long-term prediction gain value, and the non-corrupted frames include second long-term prediction lag values and second long-term prediction gain values, and wherein the second long-term prediction lag values include a last long-term prediction lag value and the second long-term prediction gain values include a last long-term prediction gain value, and the speech sequences include stationary and non-stationary speech sequences, and a first signal is used to indicate the corrupted frame, said system comprising:

a first means, responsive to the first signal, for determining whether the speech sequence in which the corrupted frame is arranged is stationary or non-stationary, and for

providing a second signal indicative of said determining;

a second means, responsive to the second signal, for replacing the first long-term prediction lag value in the corrupted frame with the last long-term prediction lag value when said speech sequence is stationary, and replacing the first long-term prediction lag value in the corrupted frame with a third lag value when said speech sequence is non-stationary.

12. The system of claim 11, wherein the third lag value is determined based on the second long-term prediction lag values and an adaptively-limited random lag jitter.

13. The system of claim 11, wherein the second means further replaces the first long-term prediction gain value in the corrupted frame with a third gain value when said speech sequence is non-stationary.

14. The system of claim 13, wherein the third gain value is determined based on the second long-term prediction gain values and an adaptively-limited random gain jitter.

15. The system of claim 11, wherein the stationary speech sequences include voiced sequences, and the non-stationary speech sequences include unvoiced sequences.

16. A decoder for synthesizing speech from an encoded bit stream, wherein the encoded bit stream includes a plurality of speech frames arranged in speech sequences, and the speech frames include at least one corrupted frame preceded by one or more non-corrupted frames, wherein the corrupted frame includes a first long-term prediction lag value and a first long-term prediction gain value, and the non-corrupted frames include second long-term prediction lag values and second long-term prediction gain values, and wherein the second long-term prediction lag values include a last long-term prediction lag value and the second long-term prediction gain values include a last long-term prediction gain value and the speech sequences include stationary and non-stationary speech sequences, and a first signal is used to indicate the corrupted frame, said decoder comprising:

a first means, responsive to the first signal, for determining whether the speech

sequence in which the corrupted frame is arranged is stationary or non-stationary, and for providing a second signal indicative of said determining;

a second means, responsive to the second signal, for replacing the first long-term prediction lag value in the corrupted frame with the last long-term prediction lag value when said speech sequence is stationary, and replacing the first long-term prediction lag value in the corrupted frame with a third lag value when said speech sequence is non-stationary.

17. The decoder of claim 16, wherein the lag value is determined based on the second long-term prediction lag values and an adaptively-limited random lag jitter.

18. The decoder of claim 16, wherein the second means further replaces the first long-term gain value in the corrupted frame with a third gain value when said speech sequence is non-stationary.

19. The decoder of claim 18, wherein the third gain value is determined based on the second long-term prediction gain values and an adaptively-limited random gain jitter.

20. The decoder of claim 16, wherein the stationary speech sequences include voiced sequences, and the non-stationary speech sequences include unvoiced sequences.

21. A mobile station, which is arranged to receive an encoded bit stream containing speech data indicative of speech signals, wherein the encoded bit stream includes a plurality of speech frames arranged in speech sequences, and the speech frames include at least one corrupted frame preceded by one or more non-corrupted frames, wherein the corrupted frame includes a first long-term prediction lag value and a first long-term prediction gain value, and the non-corrupted frames include second long-term prediction lag values and second long-term prediction gain values, and wherein the second long-term prediction lag values include a last long-term prediction lag value and the second long-term prediction gain values include a last long-term prediction gain value and the speech sequences include stationary and non-stationary speech sequences, and wherein a first signal is used to indicate the corrupted

frame, said mobile station comprising:

a first means, responsive to the first signal, for determining whether the speech sequence in which the corrupted frame is arranged is stationary or non-stationary, and for providing a second signal indicative of said determining; and

a second means, responsive to the second signal, for replacing the first long-term prediction lag value in the corrupted frame with the last long-term prediction lag value when said speech sequence is stationary, and replacing the first long-term prediction lag value in the corrupted frame with a third lag value when said speech sequence is non-stationary.

22. The mobile station of claim 21, wherein the third lag value is determined based on the second long-term prediction lag values and an adaptively-limited random lag jitter.

23. The mobile station of claim 21, wherein the second means further replaces the first long-term gain value in the corrupted frame with a third gain value when said speech sequence is non-stationary.

24. The mobile station of claim 23, wherein the third gain value is determined based on the second long-term prediction gain values and an adaptively-limited random gain jitter.

25. The mobile station of claim 21, wherein the stationary speech sequences include voiced sequences, and the non-stationary speech sequences include unvoiced sequences.

26. An element in a telecommunication network, which is arranged to receive an encoded bit stream containing speech data from a mobile station, wherein the speech data includes a plurality of speech frames arranged in speech sequences, and the speech frames include at least one corrupted frame preceded by one or more non-corrupted frames, wherein the corrupted frame includes a first long-term prediction lag value and a first long-term prediction gain value, and the non-corrupted frames include second long-term prediction lag values and second long-term prediction gain values, and wherein the second long-term prediction lag values include a last long-term prediction lag value and the second long-term prediction gain

values include a last long-term prediction gain value and the speech sequences include stationary and non-stationary speech sequences, and wherein a first signal is used to indicate the corrupted frame, said element comprising:

5 a first means, responsive to the first signal, for determining whether the speech sequence in which the corrupted frame is arranged is stationary or non-stationary, and for providing a second signal indicative of said determining; and

10 a second means, responsive to the second signal, for replacing the first long-term prediction lag value in the corrupted frame with the last long-term prediction lag value when said speech sequence is stationary, and replacing the first long-term prediction lag value in the corrupted frame with a third lag value when said speech sequence is non-stationary.

27. The element of claim 26, wherein the third long-term prediction lag value is determined based on the second long-term prediction lag values and an adaptively-limited random lag jitter.

15 28. The element of claim 26, wherein the third means further replaces the first long-term prediction gain value with a third gain value when said speech sequence is non-stationary.

20 29. The element of claim 28, wherein the third gain value is determined based on the second long-term prediction gain values and an adaptively-limited random gain jitter.

30. The element of claim 26, wherein the stationary speech sequences include voiced sequences, and the non-stationary speech sequences include unvoiced sequences.

Abstract of the Disclosure

A method and system for concealing errors in one or more bad frames in a speech sequence as part of an encoded bit stream received in a decoder. When the speech sequence is voiced, the LTP-parameters in the bad frames are replaced by the corresponding parameters in the last frame. When the speech sequence is unvoiced, the LTP-parameters in the bad frames are replaced by values calculated based on the LTP history along with an adaptively-limited random term.

09702540-103100



FIG. 1



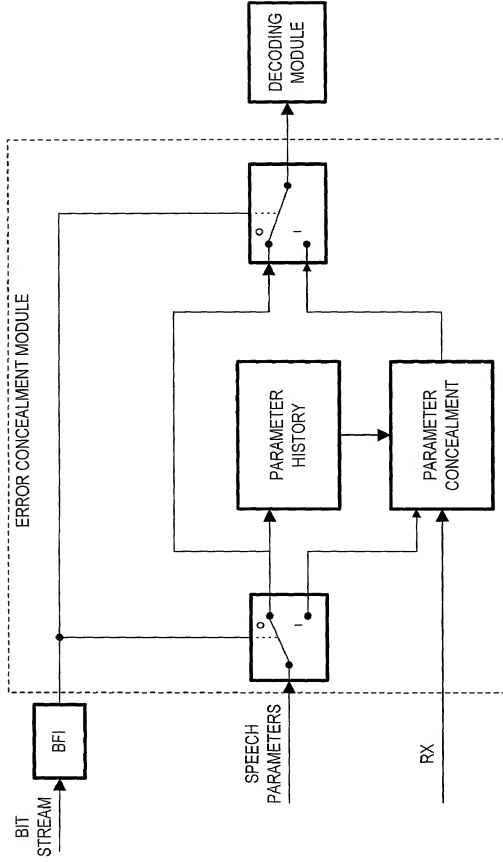


FIG. 2

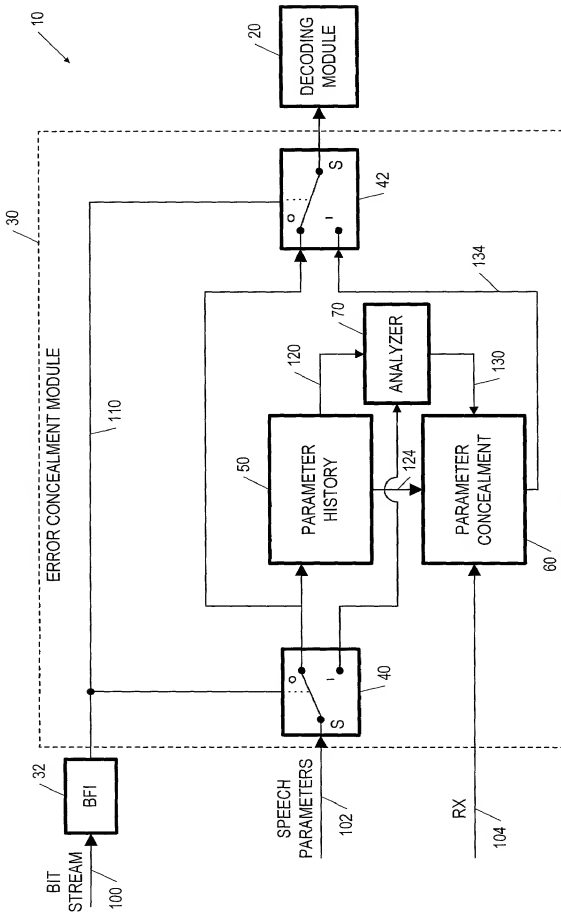


FIG. 3

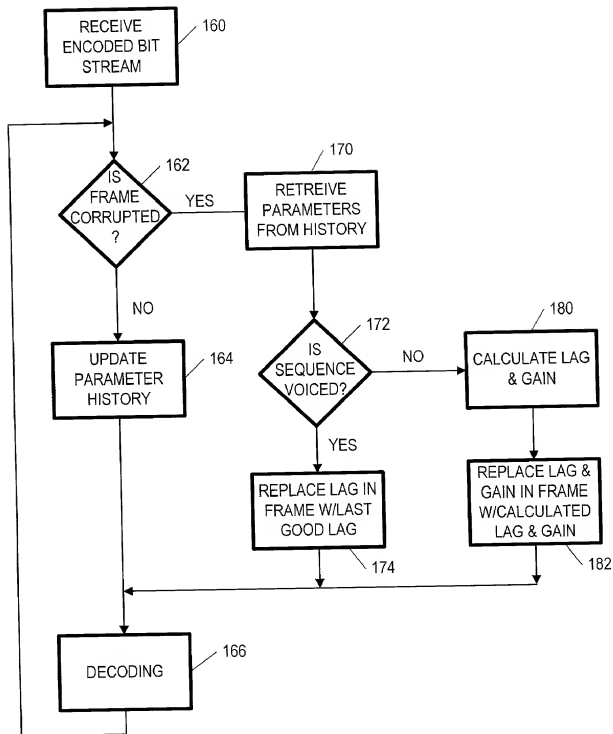


FIG. 4

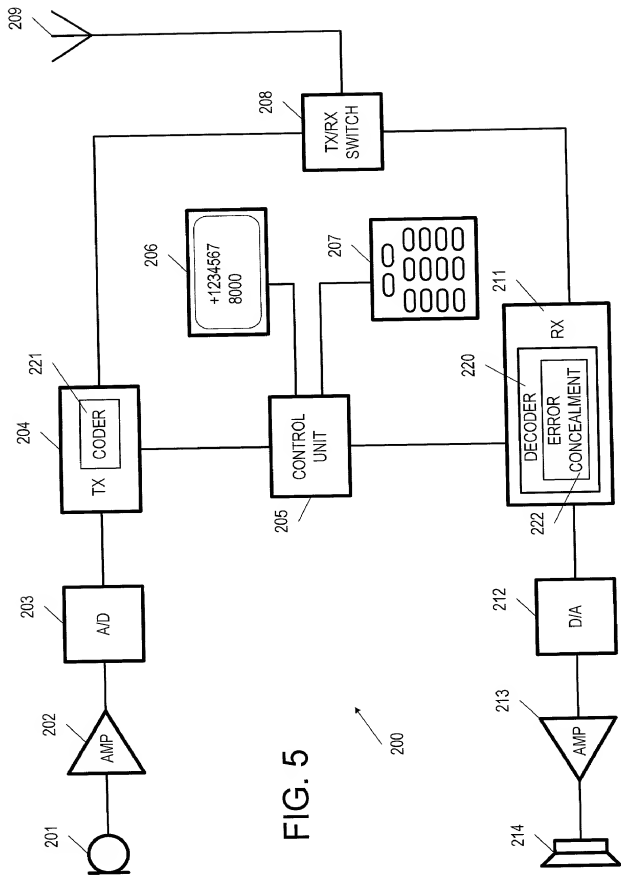


FIG. 5

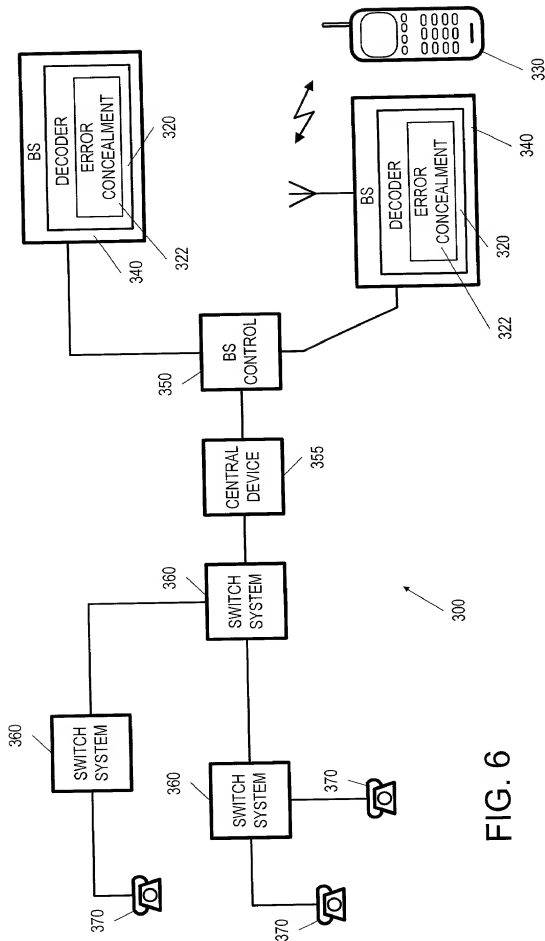


FIG. 6

001201-04520260

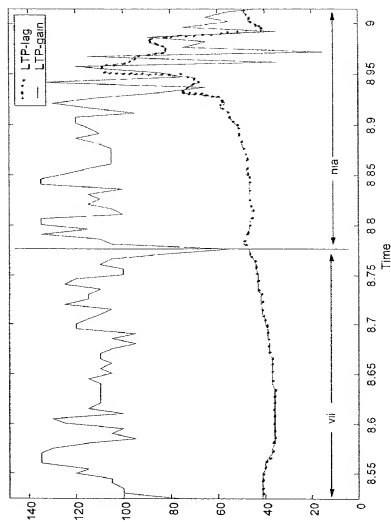


FIG. 7

007E01" 04520760

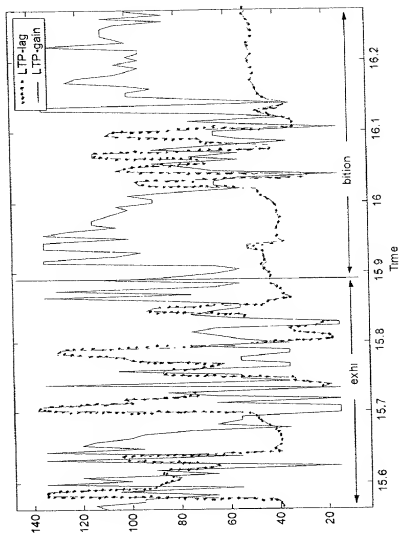


FIG. 8

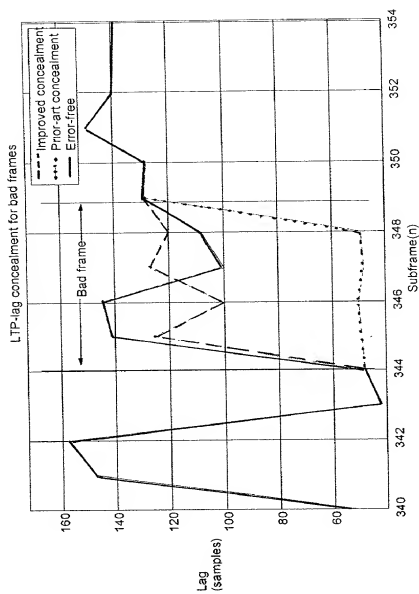


FIG. 9



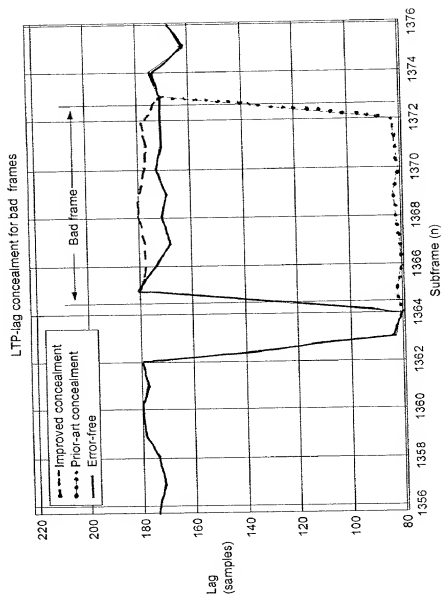
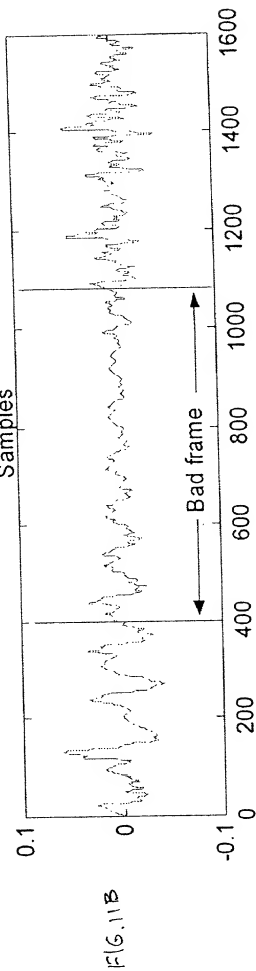
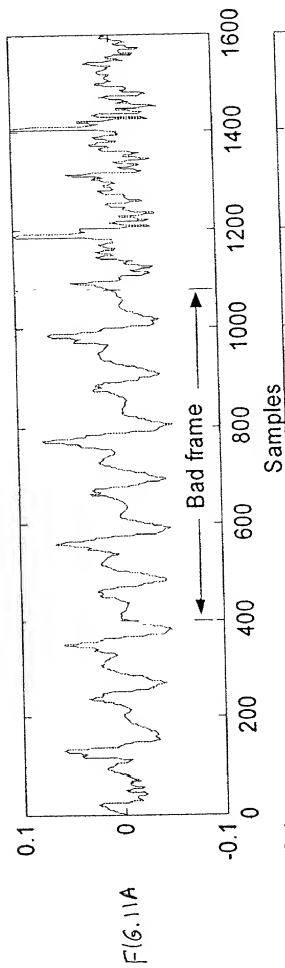


FIG. 10



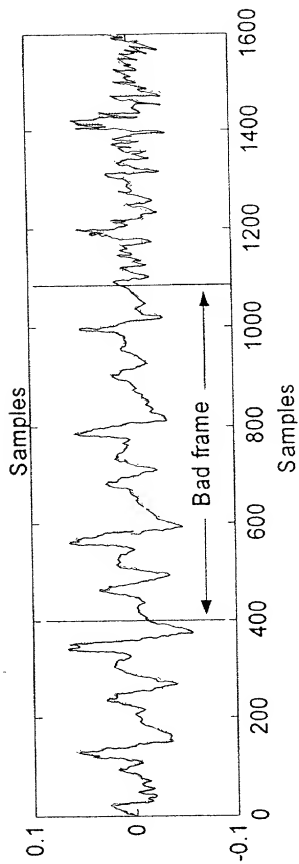


FIG. 11C